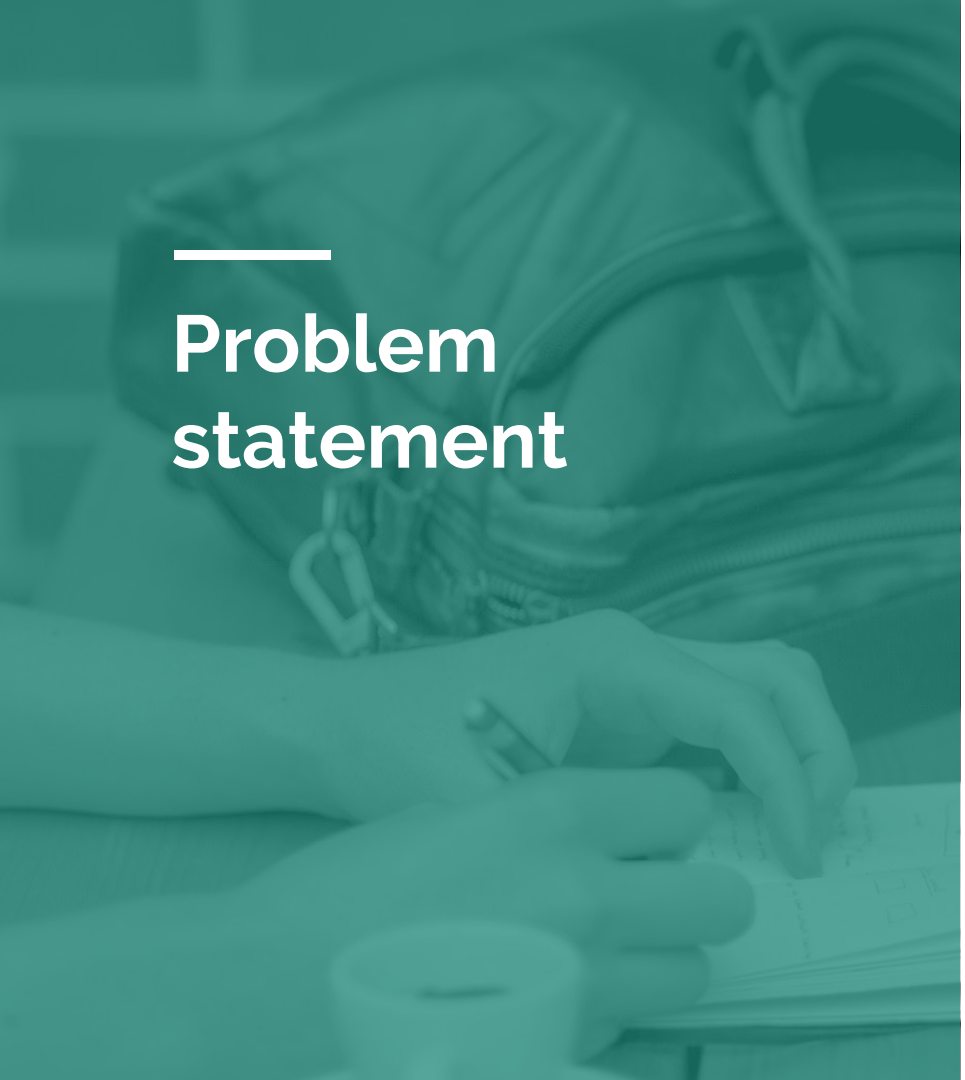
Creating Music with AI

“Hey alexa. Write a song for me.”

Outline



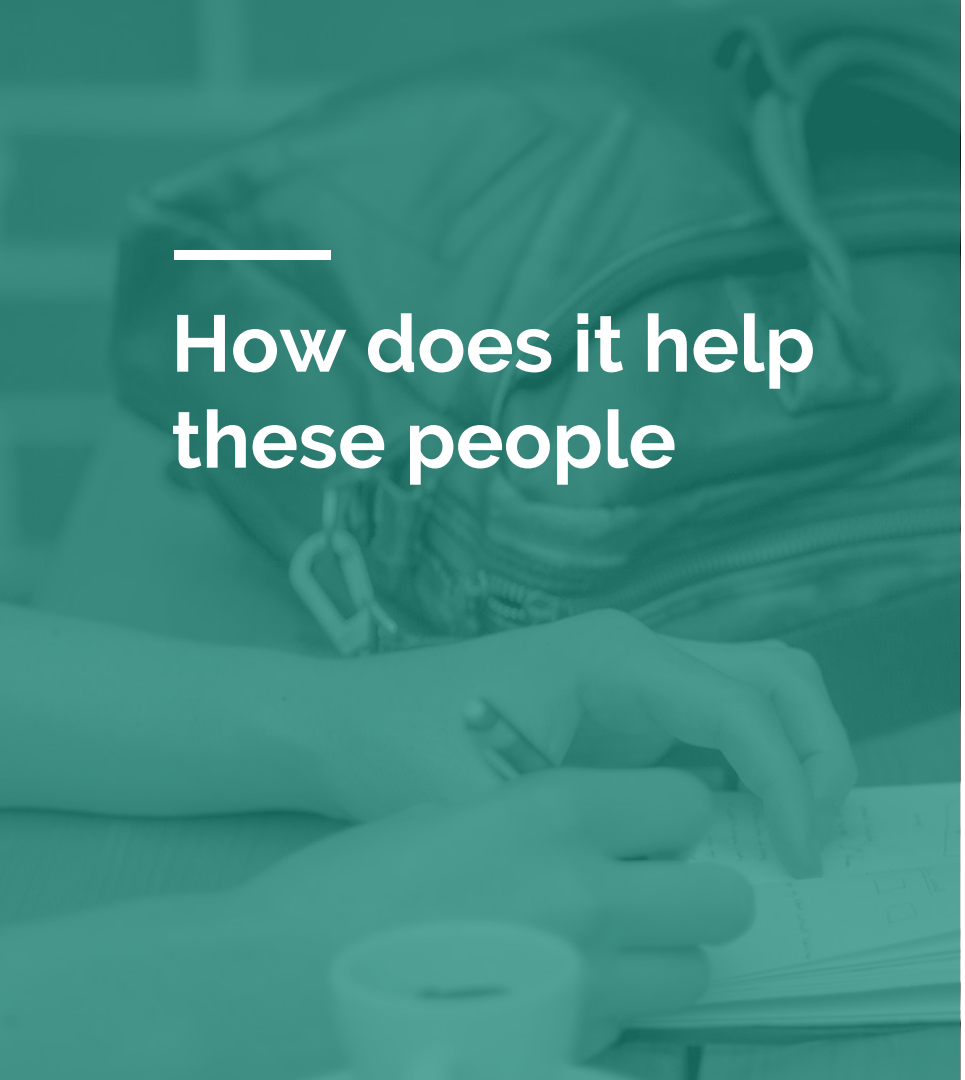
Problem statement

- Create an AI-based system to generate elements of music - Lyrics, Instrumentals, Vocals, Audio mixing.

- Motivation - Kenny Sebastian’s on-the-spot song segment in his live shows.

Who is it for?

- Music Composers / songwriters
- Poets /lyricists
- Anybody looking to be entertained.
- People trying to break into music but lack the “intuition” to compose.
- People passionate for technology
- People passionate for music



How does it help these people

- Getting over composer's block/writer's block
- Suggestive lyrics from a stream of words and from a set of words
- Suggestive music (piano, violin)
- AI - based composition (combining the two)
- Gold Standard - A system that given some words and some music, or either of them or none of them (input would be a genre or a mood), comes up with a complete musical composition.

Competition

- No existing application that reliably delivers this tech. (We could be the first) .
- Amper - Proprietary library of live-recorded instrument samples and musician-trained Creative AI
- Magenta’s NSynth - Provides artists with intuitive control over timbre and dynamic.
- Watson Beat - Human assisted music samples generation with Watson

Market research

- Massive space in market for this kind of product.
- Even if a product comes out in near future, if developed in west, it'll be built keeping in mind, western music; hence still leaving huge space in Indian/Oriental music industry.

Our current solution

• A lyrics generation system using LSTM. We trained it on all Pink Floyd songs lyrics. Hence the generated lyrics retain that flavour

• There are two modes of lyrics generation - Stream of words (denoting beginning of lyrics) and set of upto 5 words (denoting presence in lyrics).

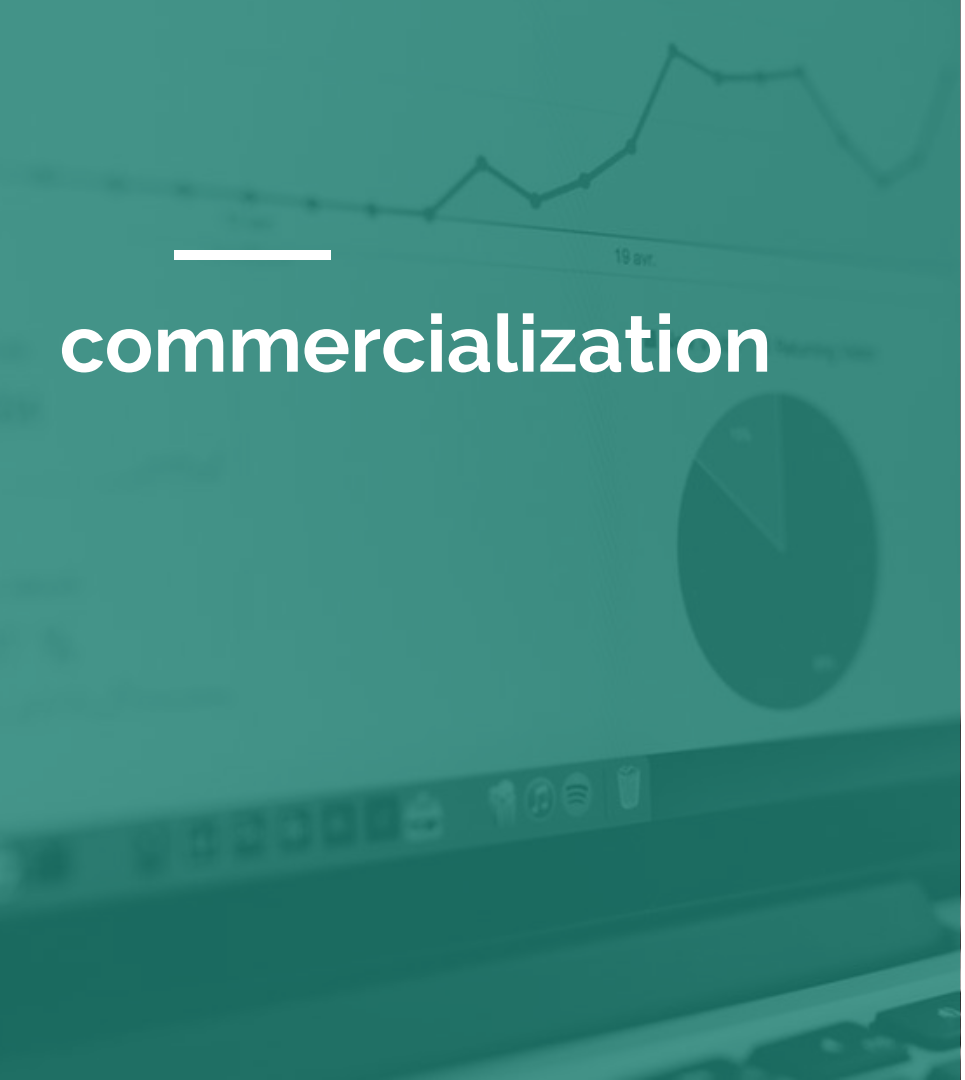
• AI powered piano - Given a piano piece extends that piece in Mozart style. Model trained using Seq2Seq model on midi files of Mozart's symphonies.

Modelling Challenges

- Representing music with numbers effectively is tough mostly because songs are information-wise very rich (there are multiple components at work which make a song pleasant).
- There is always a problem of balance to be maintained between non-trivial content and repetitive content

What might lie ahead

- Lyrics generation system can be made more robust simply by training on more data.
- We can have saved model checkpoints trained genre + artist + mood wise, so that as per user query the suitable checkpoint will do the inference using 1MSong dataset.
- More alternatives for AI generated piano piece can be explored including GANs.



commercialization

- Per API call pricing model with few initial calls free
- Subscriber system can only be explored once the product's performance is superior.
- Notice that every competitor we saw were premium services mostly coming from big giants.
- Also region-wise attractive pricing can be very helpful in initial stage because the competitor is Google.

Thank You.

Any questions?

References

1. Research from OpenAI (2017) by Christine McLeavey Payne:

http://christinemcleavey.com/clara-a-neural-net-music-generator/

2. An End-to-End Neural Network for Polyphonic Piano Music Transcription, Sigtia et al.

https://arxiv.org/pdf/1508.01774.pdf

Appendix - Implementation details

What goes inside model - 49 dimensional vectors of 0s and 1s, each representing 0.02 seconds of sound.

Each of the 49 numbers (0 or 1) denote the presence of that note at that time.

Not surprisingly, each instrument will ideally have a its own input shape for the model depending on how many configurations are there, thus making combining instruments extremely tough with this approach. The obvious solution is to feed music pieces instead of instrumentals into model. But then the model doesn't learn anything (**Each instrument has a different latent distribution of frequencies/timbre/other stuff that make music 'music', so it would be mathematically imprecise to have a vector represent all instruments' sounds**)

Competition

- No existing application that reliably delivers this tech. (We could be the first) .
- Amper - It builds tracks from pre-recorded samples and spits out actual audio, not MIDI. From there, you have to change the tempo, the key; mute individual instruments, or switch out entire instrument kits to shift the mood of the song it's made.
- Three biggest competitors: IBM Watson Beat, Google Magenta's NSynth Super and Amper Music
- On a micro scale, the music is convincing, but the longer you listen, the less sense it makes. None of them are good enough to craft a Grammy Award-winning song on their own yet.